Higher-order correlations in common input shapes the output spiking activity of a neural population

Lisandro Montangie, Fernando Montani *

IFLYSIB, CONICET & Universidad Nacional de La Plata, Calle 59-789, (1900) La Plata, Argentina
Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Calle 49 y 115. C.C. 67, (1900) La Plata, Argentina

HIGHLIGHTS

- Complexity of neuronal synchronous activity patterns.
- To capture subtle changes in the common neuronal inputs.
- Input statistics transformed into higher-order output statistics.
- Neuronal inputs and the spiking outputs follow q-Gaussian statistics.
- Importance of higher-order correlations at population level.

ARTICLE INFO

ABSTRACT

Recent neurophysiological experiments suggest that populations of neurons use a computational scheme in which spike timing is regulated by common non-Gaussian inputs across neurons. The presence of beyond-pairwise correlations in the neuronal inputs and the spiking outputs following a non-Gaussian statistics elicits the need of developing a new theoretical framework taking into account the complexity of synchronous activity patterns. To this end, we quantify the amount of higher-order correlations in the common neuronal inputs and outputs of a population of neurons. We provide a novel formalism, of easy numerical implementation, that can capture the subtle changes of the inputs heterogeneities. Within our approach, correlations across neurons arise from q-Gaussian inputs into threshold neurons and higher-order correlations in the spiking outputs activity are quantified by the parameter q. We present an exhaustive analysis of how input statistics are transformed in this threshold process into output statistics, and we show under which conditions higher-order correlations can lead to either bigger or smaller number of synchronized spikes in the neural population outputs.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Neurons primarily communicate with each other by generating sequences of action potentials. The spiking activity of nearby cortical neurons is not independent and several studies have explored the importance of this correlated activity. A great deal of attention has been devoted to answer the question of whether the information conveyed by the activity of an ensemble of neurons is determined solely by the number of action potentials fired by each cell independently or correlations across neurons may be also important for information transmission [1–17]. It has been proposed that pairwise correlations

are important for information representation or processing in retina [18,19], thalamus [20] and cerebral [10,11,21] and cerebellar cortices [22,23]. However, recent studies suggest that higher-order correlation structures are also quite crucial to get a better understanding of how information might be transmitted in the brain [18,19,24–35]. In this framework, pairwise modeling does not provide a fair description of the overall organization of neural interactions when considering neuronal systems in general [18,19,24–35].

Several studies have pointed out the relevance of higher-order correlations (HOCs), as pairwise models do not explain the variability in activity patterns at a more general level [27–34,36]. Indeed neurophysiological research has shown that pairwise models fail to explain the responses of spatially localized triplets of cells [29–31,37] when describing the activity of large neuronal populations responding to natural stimuli [31]. Deviations from the Pairwise Maximum Entropy (PME) model indicate that HOCs have to be taken into account for modeling population statistics [24,38–40]. Thus, the intricacy of the neurophysiological data highlights the need to develop a theoretical framework accounting for the statistical complexity of synchronous activity patterns. Pattern probabilities for the so-called Dichotomized Gaussian (DG) model [24,37–41] were developed using the cumulative distribution of multivariate Gaussians showing high precision fitting of the experimental data, and therefore evidencing that HOCs are required to properly account for cortical dynamics. More specifically, in the DG model binary patterns are generated by thresholding a multivariate Gaussian random variable, and correlations between neurons arise from correlations in the underlying Dichotomized Gaussian distribution [24,37–41]. The DG model can be interpreted as a phenomenological model of how pairwise correlations in the inputs to threshold neurons shape the distribution of spike-counts in the population [42]. It has been extensively used to construct quantitative predictions on how departures from pairwise models depend on common Gaussian like neuronal inputs [40] and to generate population spike trains with specified mean and pairwise statistics [39,43] (via sampling from the latent Gaussian). However, it is important to point out that the DG model does not explicitly quantify what is the amount of HOCs in the spiking population output. It just produces HOCs in the spiking neuronal outputs that cannot be obtained by using pure pairwise models [38–40].

Neurophysiology research using whole-cell patch-clamp recordings in vivo has been used to measure subthreshold membrane potential fluctuations, showing evidence of non-Gaussian presynaptic input distributions in the olfactory bulb of the drosophila and the auditory cortex of rats [44,45]. Importantly, the existence of non-Gaussian membrane potential dynamics implies sparse synchronous activity in the auditory cortex of rats and in presynaptic olfactory receptor neurons of Drosophila. The subthreshold membrane potentials on single trials have been analyzed making inferences about the underlying population network activity, finding that during both spontaneous and evoked responses it was highly non-Gaussian. These dynamics suggest a computational scheme in which spike timing is controlled by common non-Gaussian input across neurons [44,45]. The intricacy of the neuronal networks suggests therefore the need of taking into consideration the non-Gaussian input connectivity across neuronal inputs. Thus, to gain a better understanding of how neural information is processed in the brain, we need to account for HOCs in the input connectivity across neuronal inputs and spiking outputs. To this end, we consider the q-Gaussian distribution [46], which is a generalization of the Gaussian distribution in the same way that Tsallis entropy is a generalization of standard Boltzmann–Gibbs entropy [47]. This distribution is characterized by a parameter q, which conveys departure from pure pairwise correlations.

In this paper, we statistically characterize the population firing activity obtained from simultaneous recordings of neurons across all layers of a simulated cortical microcolumn. We put special emphasis on how to perform computational estimations efficiently when considering HOCs in the input connectivity, providing an appropriate mathematical framework to quantify HOCs in the spiking output of the neuronal population. We take advantage of the fact that the q-Gaussian distribution, in the heavy tail region, is equivalent to the Student's t-distribution with a direct mapping between the deformation parameter q and the degree of freedom of the Student's t-distribution. This allows us to write the joint probability distribution of firing of the population of neurons in terms of the Student's t-distribution cumulative distribution function accounting for beyond-pairwise correlations in their inputs, and, more importantly, to provide a robust estimation of output statistics in terms of bivariate q-Gaussian distributions. Three types of models are considered: an independent model which shows no correlations across spikes, a DG model that allows to generate HOCs in the output by considering pairwise-correlated inputs, and a Dichotomized non-Gaussian model (q-DG) which accounts for HOCs in both the input and output statistics. This scheme highlights the importance of both non-Gaussian common inputs and of the HOCs within the output statistics, when describing the structure of cortical networks.

2. Methodology

2.1. The ECLT framework

Although inputs in the DG model are Gaussian distributed and, therefore, there are no interactions beyond second order in their inputs, the nonlinear threshold spiking may give rise to statistical interactions of all orders in their spiking outputs [24,37–41]. Thus, using the DG approach, common pairwise input could be used to produce HOCs in the output population activity. This method has been developed within the Central Limit Theorem (CLT) framework, which ensures that the probability distribution function of any measurable quantity is a normal Gaussian distribution, provided that a sufficiently large number of independent random variables with exactly the same mean and variance are being considered [48]. Be that as it may, the CLT does not hold if correlations between random variables cannot be neglected. However, in the presence

of weak or strong correlations of any sort, the CLT has been generalized in recent publications by M. Gell-Mann, C. Tsallis, S. Umarov, C. Vignat and A. Plastino [46,49–52]. They have shown that when a system with weakly or strongly correlated random variables is being considered, if we gather a sufficiently large number of such systems together, the probability distribution will converge to a q-Gaussian distribution [46,49–52].

In order to understand how neural systems perform computations and process sensory information, we need to understand how the network structure of firing patterns across a population of neurons is built. Information in neural populations is often encoded in the activity of large, highly interconnected ensembles, so that higher-order output statistics can also be shaped and modulated by higher-order input statistics and their intrinsic circuit mechanisms [24,29,31,37]. In order to take this into account, we have considered in Refs. [33,34] that each neuron is subject to a weighted sum of common inputs which are q-Gaussian [53] due to the ECLT [46,50], in which the independence constraint for the independent and identically distributed variables is relaxed to an extent defined by the q parameter [46,49–52,54]. In the framework of q-algebra, the corresponding generalization of the CLT becomes possible and relatively simple. When a system with weakly or strongly correlated random variables is being considered, if we gather a sufficiently large number of such systems together, the probability distribution will converge to a q-Gaussian distribution [46,49–52]. We use the "natural extension" of the CLT proposed in [50], which accounts for cases in which correlations between random variables are non-negligible. Thus, q-Gaussians are the probability density functions in the ECLT:

$$\mathcal{G}_q(x; \beta) = \frac{\sqrt{\beta}}{C_q} \exp_q\left(-\beta x^2\right) \tag{1}$$

where $\exp_q$ is defined as the q-exponential [53]

$$\exp_q(x) = \begin{cases} \exp(x) & \text{if } q = 1 \\ [1 + (1-q)x]^{\frac{1}{1-q}} & \text{if } q \neq 1 \text{ and } 1 + (1-q)x > 0 \\ 0 & \text{if } q \neq 1 \text{ and } 1 + (1-q)x \leq 0 \end{cases} \tag{2}$$

and the normalization factor C_q is given by

$$C_q = \begin{cases} \dfrac{2\sqrt{\pi}\,\Gamma\left(\frac{1}{1-q}\right)}{(3-q)\sqrt{1-q}\,\Gamma\left(\frac{3-q}{2q-2}\right)} & \text{for } -\infty < q < 1, \\ \sqrt{\pi} & \text{for } q = 1, \\ \dfrac{\sqrt{\pi}\,\Gamma\left(\frac{3-q}{2q-2}\right)}{\sqrt{1-q}\,\Gamma\left(\frac{1}{1-q}\right)} & \text{for } 1 < q < 3. \end{cases} \tag{3}$$

β is a positive real number and q is a (problem-dependent) positive real index. Notice that in the limit of $q = 1$ a normal Gaussian distribution is recovered as $\lim_{N\to\infty}(1 + \frac{1}{N})^N = e$, which can be rewritten as $\lim_{q\to 1}[1 + (1-q)]^{\frac{1}{(1-q)}} = e$. Furthermore, when $q = 2$ the q-Gaussian corresponds to a Cauchy distribution, which is also known as the Lorentz distribution.

Hence, perhaps undetectable, higher-order input correlations may well have an important effect at population level [34]. Moreover, this extension is achieved with the inclusion of just one extra parameter, the deformation parameter q, so that we remain in a low dimensional space, avoiding the sampling size problem of a complete description of HOCs [55]. Importantly, our approach converges to the DG model of Amari and collaborators when we consider the limit of the CLT framework (that is, $q \to 1$).

Furthermore, the q-Gaussian distribution is often favored for its heavy tails in comparison to the Gaussian for $1 < q < 3$. In the heavy tail region, the distribution is equivalent to the Student's t-distribution (see Eq. (4)) with a direct mapping between q and the degrees of freedom ν

$$\mathcal{T}_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \tag{4}$$

Student's t-distribution is a well known and commonly used distribution in statistics, quantifying deviations from the mean in samples for which the standard deviation is unknown. In this sense, statistically it is a scaled reparametrization of the Student's t-distribution to describe small-sample statistics, where the parameters q and β are related to ν. Given a Student's t distribution with ν degrees of freedom, the equivalent q-Gaussian has

$$q = \frac{\nu + 3}{\nu + 1} \quad \text{with} \quad \beta = \frac{1}{3 - q} \tag{5}$$

and inverse

$$\nu = \frac{3 - q}{q - 1}, \tag{6}$$

but only if $\beta = \frac{1}{3-q}$. Whenever $\beta \neq \frac{1}{3-q}$, the function is simply a scaled version of Student's t distribution. This transformation is important for its sampling properties and availability of software packages.

In the case of bivariate q-Gaussian distributions [50], the definition is

$$\mathcal{G}_q(\mathbf{x}; \rho) = \frac{1}{K_q\sqrt{1-\rho^2}} \, \exp_q\left(-\frac{\|\mathbf{x}\|^2}{1-\rho^2}\right) \tag{7}$$

for $\mathbf{x}$ real 2-dimensional column vectors, where ρ is the pairwise correlation and the normalization factor K_q is given by

$$K_q = \begin{cases} \dfrac{\pi(6-4q)\Gamma\left(\frac{2-q}{1-q}\right)}{(1-q)\Gamma\left(\frac{2-q}{1-q}+1\right)} & \text{for } -\infty < q < 1 \\[4mm] \pi & \text{for } q = 1 \\[2mm] \dfrac{\pi(6-4q)\Gamma\left(\frac{1}{q-1}-1\right)}{(1-q)\Gamma\left(\frac{1}{q-1}\right)} & \text{for } 1 < q < \dfrac{3}{2}. \end{cases} \tag{8}$$

It should be considered that this bivariate distribution further restricts the possible values of the deformation parameter q.

2.2. A simple application of the ECLT within the information geometry framework

In the information geometry framework [55,56], neuronal firing in a population of size N is represented as a binary vector $\mathbf{x} = (x_1, \ldots, x_N)$, where $x_i = 0$ if neuron i is silent in some time window ΔT and $x_i = 1$ if it is firing. Then, for a given time window, we consider the set the probability distribution of binary vectors, $\{p(\mathbf{x})\}$, which consists of 2^N probabilities

$$p(\mathbf{x}) = \text{Prob}\{x_1 = i_1, \ldots, x_N = i_N\} = p_{i_1 \ldots i_N} \tag{9}$$

subject to the normalization

$$\sum_{i_1, \ldots, i_N = 0, 1} p_{i_1 \ldots i_N} = 1. \tag{10}$$

As proposed by Amari and co-workers (see, for instance, Refs. [55,56]), the set of all the probability distributions $\{p(\mathbf{x})\}$ forms a $(2^N - 1)$-dimensional manifold $\mathbf{S}_N$. This approach uses the orthogonality of the natural and expectation parameters in the exponential family of distributions. It is also useful for analyzing neural firing in a systematic manner based on information geometry. Any such probability distribution can be unequivocally determined using a "coordinate system". One possible coordinate system is given by the set of $2^N - 1$ marginal probability values:

$$\eta_i = E[x_i] = \text{Prob}\{x_i = 1\}, \quad i = 1, \ldots, N \tag{11}$$

$$\eta_{ij} = E[x_i x_j] = \text{Prob}\{x_i = x_j = 1\}, \quad i < j \tag{12}$$

$$\vdots$$

$$\eta_{123\ldots N} = E[x_1 \ldots x_N] = \text{Prob}\{x_1 = x_2 = \cdots = x_N = 1\}. \tag{13}$$

These are called the $\boldsymbol{\eta}$-coordinates [56]. Moreover, provided $p(\mathbf{x}) \neq 0$, any such distribution can be expanded as in Ref. [55]

$$\log p(\mathbf{x}) = \sum_{i=1}^{N} x_i\,\theta_i + \sum_{i<j} x_i x_j\,\theta_{ij} + \sum_{i<j<k} x_i x_j x_k \theta_{ijk} + \sum_{i<j<k<l} x_i x_j x_k x_l\,\theta_{ijkl} + \cdots + x_1 \cdots x_N\,\theta_{1\ldots N} - \psi, \tag{14}$$

where there are in total $2^N - 1$ different θ correlation coefficients that can be used to determine univocally the probability distribution. The $\boldsymbol{\theta}$ forms a coordinate system, named $\boldsymbol{\theta}$-coordinates.

In the following paragraphs we provide an intuitive explanation the connection between ECLT with the information geometry framework. Let us now consider a q-Gaussian distribution of x, with mean μ and variance σ. Note that in the literature, the latter can be also found as q-mean and q-variance. ψ denotes a normalization constant.

$$\mathcal{G}_q(x; \mu, \sigma) = \exp_q\left[-\beta\left(\frac{x-\mu}{\sigma}\right)^2 - \psi\right]$$

$$= \left\{1 + (q-1)\left[\beta\left(\frac{x-\mu}{\sigma}\right)^2 - \psi\right]\right\}^{\frac{1}{1-q}}, \tag{15}$$

which can be rewritten as

$$
\mathcal{G}_q(x; \mu, \sigma) = \exp\left\{\frac{1}{1-q} \ln\left[1 + (q-1)\left(\beta\left[\frac{x-\mu}{\sigma}\right]^2 - \psi\right)\right]\right\} \tag{16}
$$

and using the definition for the logarithmic function

$$
\ln(1+x) = -\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \tag{17}
$$

then

$$
\mathcal{G}_q(x; \mu, \sigma) = \exp\left\{-\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(q-1)^{n-1}\left[\beta\left(\frac{x-\mu}{\sigma}\right)^2 - \psi\right]^n\right\}
$$
$$
= \exp\left[-\beta\left(\frac{x-\mu}{\sigma}\right)^2\right]\exp\left\{\psi - \sum_{n=2}^{\infty} \frac{(-1)^{n+1}}{n}(q-1)^{n-1}\left[\beta\left(\frac{x-\mu}{\sigma}\right)^2 - \psi\right]^n\right\}. \tag{18}
$$

This expression can be rearranged as

$$
\mathcal{G}_q(x; \mu, \sigma) = \exp\left\{-\left[\beta\left(\frac{x-\mu}{\sigma}\right)^2\right] + \tilde{\Psi}(x, q)\right\}, \tag{19}
$$

where

$$
\tilde{\Psi}(x, q) = \psi - \sum_{n=2}^{\infty} \frac{(-1)^{n+1}}{n}(q-1)^{n-1}\left[\beta\left(\frac{x-\mu}{\sigma}\right)^2 - \psi\right]^n. \tag{20}
$$

Thus, this suggests that in ECLT framework correlations are taken into account through the parameter q which acts as a "colored noise" in $\tilde{\Psi}(x, q)$.

2.3. The dichotomized q-Gaussian model

A population of N neurons, if we assume spike trains to be discretized into sufficiently small time bins such that each time bin contains at most one spike, can be represented by a collection of N binary sequences. That is, each neuron can either spike or not at any given time and thus the population can be in any of 2^N different states. Estimating the full joint distribution over these 2^N patterns becomes rapidly infeasible for increasing numbers of neurons N, due to combinatorial explosion. However, it is often possible to measure the first and second moments of the given distribution. These are the firing rates of individual cells and their pairwise correlations. The DG model allowed to generate spike trains from a population of N neurons where these moments have been specified. However, the distribution of the sum of inputs to that population was assumed to be Gaussian and thus, not to include possible beyond-pairwise correlations [38–40]. Thus, we would like to be able to relax the non-Gaussianity "constraints" in the common input distributions in order to take into account the impact of HOCs.

We model therefore a multivariate, N-dimensional binary random variable $X \in \{0, 1\}^N$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. A sample from the q-DG distribution is obtained by first drawing a sample from a N-dimensional q-Gaussian random variable U and then thresholding it into 0 and 1. That is,

$$
X_i = 1 \text{ if and only if } U_i > 0 \quad \text{where } U \sim \mathcal{G}_q(h, A), \tag{21}
$$

zero otherwise.

This thresholding operation will change the moments of the output, so X will, in general, not have the same mean and covariance as U. The moments of the output are given by the means $\mu_i = \langle X_i \rangle$ and pairwise covariances $\Sigma_{ij} = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle$. However, the effect of the truncation can be calculated [57], as is the case for the Dichotomized Gaussian model, and corrected for. Samples from X are constructed by generating samples from U (with mean h and covariance A), and thresholding it at 0, i.e. setting $X_i = 1$ if and only if $U_i > 0$.

We can choose the mean h and covariance A of U such that after truncation, X has the desired moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. For an easier algorithmic implementation, we exploit the equivalence, if conditions are met, between q-Gaussian distributions and Student's t-distributions. Assuming (without loss of generality) unit variances for U, the mean spiking probabilities $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ of X are given by

$$
\mu_i = \Phi_q(h_i), \tag{22}
$$
$$
\Sigma_{ii} = \Phi_{q,2}(h_i, -h_i, 0), \tag{23}
$$
$$
\Sigma_{ij} = \Phi_{q,2}(h_i, h_j, A_{ij}) - \Phi_{q,2}(h_i, h_i, 0). \tag{24}
$$

Here, $\Phi_q(.)$ is the cumulative distribution function of a univariate q-Gaussian with mean 0 and unit variance from now on referred to as q-Normal, and $\Phi_{q,2}(.,.;\rho)$ the cumulative distribution function of a bivariate q-Gaussian with unit variances and correlation coefficient ρ (bivariate q-Normal). Both distributions are characterized also by the deformation parameter q. Note that Eq. (24) is only valid when i is different from j, and that in this equation we have considered the approach that Σ_{ij} is the cumulative distribution function of a bivariate q-Gaussian minus cumulative distribution function when taking a null covariance matrix $A_{ij} = 0$. This allows us to quantify the amount of correlations in the outputs.

To be able to define q-Normal distributions (that is, for the variance of a q-Gaussian to be the unit) the deformation parameter q must be restricted to the range $[1, 5/3)$. In this range, they can be directly mapped to scaled Student's t-distributions $\mathcal{T}_\nu$. To have a unit variance, $\beta = \frac{1}{5-3q}$. Under this condition, the degree of freedom would still be $\nu = \frac{3-q}{q-1}$ and the scale parameter is $\zeta^2 = \frac{5-3q}{3-q}$, such that

$$\mathcal{N}_q(x) = \mathcal{G}_q\left(x; \frac{1}{5-3q}\right) = \frac{1}{\zeta}\mathcal{T}_\nu\left(\frac{x}{\zeta}\right)$$

$$\equiv \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\zeta^2\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left[1 - \frac{1}{\nu}\left(\frac{x}{\zeta}\right)^2\right]^{-\frac{\nu+1}{2}}. \tag{25}$$

In the next section we show how to perform the estimations of the input and output statistics of a population of neurons.

2.4. Entropy and Kullback–Leibler divergence

In information theory [58], entropy is a quantity that measures how interesting a set of responses is. Given a distribution of neural responses, if each one of those responses is identical, or if only a few different responses appear, the data set is relatively uninteresting. Computing entropy then gives a sense of variability in the firing rate distribution. In general, the entropy of a random variable X following a probability density function $p(x)$ is defined as

$$S(X) = -\sum p(x)\log p(x). \tag{26}$$

Since we are interested in quantifying entropy differences [40], we define a normalized differential of entropy ΔS_q in terms of the deformation parameter q as

$$\Delta S_q(X) = \frac{S_q(X) - S_1(X)}{S_1(X)}. \tag{27}$$

That is, it is the difference of the entropy value for a given q and the entropy value in the case of $q = 1$, normalized to the latter.

One method available to test for the presence of HOCs is by measuring the distribution of population activity in multi-unit recordings and fitting the deformation parameter q of such distribution, which represents the amount of HOCs present in the distribution of firing. To this end, we consider the Kullback–Leibler divergence (KL-divergence) [59], which constitutes an Information Theory quantity and is a measure of the "distance" between two probability distributions:

$$D_{KL}(p\|p_0) = \sum p(x)\log\frac{p(x)}{p_0(x)}. \tag{28}$$

In words, it is the expectation of the logarithmic difference between the probabilities p and p_0, where the expectation is taken using the probabilities p. Another interpretation is of the difference between two entropies, as it is also regarded as a relative entropy. It is not a true metric since it does not obey the triangle inequality, and in general $D_{KL}(p\|p_0)$ does not equal $D_{KL}(p_0\|p)$. It gives a sense of the inefficiency of assuming that the distribution is p_0 when the true distribution is p. Thus, using the KL-divergence allows us to quantify how statistically different from the $q = 1$ (that is, Gaussian) case a given measured distribution is.

In this paper, we used a normalized version of this quantity. The normalized KL-divergence is defined as the difference between the KL-divergence of a distribution f_q for a given pairwise correlation and the pairwise independent case for a given value for the deformation parameter q and the KL-divergence of a distribution f_1 of such pairwise correlation and the independent case for $q = 1$. That is,

$$\Delta_q = D_{KL}\left(f_q(\rho)\|f_q(0)\right) - D_{KL}\left(f_1(\rho)\|f_1(0)\right). \tag{29}$$

3. Results

3.1. Linking input and output HOCs

As we have discussed in the previous sections, we consider that correlations across neurons arise from q-Gaussian inputs into threshold neurons. That is, we develop in a following a non-Gaussian Dichotomized model which accounts for the

HOCs in the inputs and outputs of a neuronal population. Eq. (24) provides the relationship between the correlations of the non-Gaussian input and the correlations of the binary threshold neurons. The mean values h_i can be found by inverting the Eq. (22). That is, we can calculate these quantities using

$$h_i = \Phi_q^{-1}(\mu_i), \tag{30}$$

since it can be shown that $\Phi_q(.)$ is invertible in the range considered.

Determining Σ_{ij} can be achieved by finding a suitable value such that

$$\Sigma_{ij} - \left[\Phi_{q,2}(h_i, h_j, A_{ij}) - \Phi_{q,2}(h_i, h_j, 0) \right] = 0. \tag{31}$$

Despite this apparent simplicity, a careful analysis of the range of possible covariance matrices is needed. When dealing with simulated random variables with covariance matrices that are not directly estimated from experimental data but constructed by other considerations, not every positive definite symmetric matrix can be used as the covariance matrix of a multivariate binary distribution. This is a general property of the binary distributions and not a weakness of the particular model. Thus, there is not straightforward rule for determining the suitability of any given covariance matrix. In comparison to the Gaussian distribution, the independent bivariate q-Gaussian cumulative distribution function is not the product of individual cumulative distribution functions. Thus, feasible covariances are not only bounded by the value of means of the output but also by the deformation parameter of the input. For example, for two binary random variables X_1 and X_2 with means μ_1 and μ_2, the covariance is then bounded by the values [39]

$$\max \left\{ -\mu_{1,2}, \mu_1 + \mu_2 - (1 + \mu_{1,2}) \right\} \leq \mathrm{Cov}(X_1, X_2), \tag{32}$$

$$\mathrm{Cov}(X_1, X_2) \leq \min \left\{ \mu_1 - \mu_{1,2}, \mu_2 - \mu_{1,2} \right\} \tag{33}$$

where $\mu_{1,2} = \Phi_q([\Phi_q^{-1}(\mu_1), \Phi_q^{-1}(\mu_2)], 0)$. This provides a different set of constraints for the covariance. To show that Eq. (24) has a solution for any μ_i, μ_j and $\Sigma_{ij} = \mathrm{Cov}(X_i, X_j)$ that satisfies the conditions above, we need to show that we can find an α such that

$$\Phi_{q,2}(h_i, h_j, \alpha) = \mathrm{Cov}_{ij} + \Phi_{q,2}(h_i, h_i, 0). \tag{34}$$

As $\Phi_{q,2}$ is continuous in α, and it satisfies that

$$\Phi_{q,2}(h_i, h_j, -1) \leq \mathrm{Cov}_{ij} + \Phi_{q,2}(h_i, h_i, 0) \leq \Phi_{q,2}(h_i, h_j, 1) \tag{35}$$

we have well established the bounds of the cumulative distribution function of the bivariate q-Gaussian. In the next section, we take advantage of recent mathematical progress on q-geometry to show how to estimate $\Phi_q([a, b], \rho)$ in the current framework.

3.2. Estimations of the bivariate q-Gaussian cumulative distribution function

The standard bivariate q-Gaussian cumulative distribution function is defined by

$$\Phi_q([a, b], \rho) = \frac{\Gamma\left(\frac{1}{q-1}\right)(q-1)}{\pi \Gamma\left(\frac{1}{q-1} - 1\right)(6 - 4q)\sqrt{1 - \rho^2}} \int_{-\infty}^{a} \int_{-\infty}^{b} dy\, dx \left[1 + \left(\frac{q-1}{6-4q}\right) \frac{x^2 + y^2 - 2\rho xy}{1 - \rho^2} \right]^{\frac{1}{1-q}}, \tag{36}$$

where $a, b \in \mathbb{R}$ and $q \in [1, 3/2)$. In comparison to the univariate q-Gaussian distribution, its bivariate counterpart cannot be easily mapped into a Student's t-distribution. Furthermore, this integral cannot be solved analytically. Thus, since this function is not included in common numerical packages, we need to simplify this expression to be able to obtain accurate numerical values.

First, we use the so-called Hilhorst transform [53,60] to rewrite Eq. (36) as

$$\Phi_q([a, b], \rho) = \frac{\Gamma\left(\frac{1}{q-1}\right)(q-1)}{\pi \Gamma\left(\frac{1}{q-1} - 1\right)(6 - 4q)\sqrt{1 - \rho^2}} \int_0^{\infty} ds\, s \left(\frac{s^2}{2}\right)^{\frac{1}{q-1}} e^{-\frac{s^2}{2}}$$

$$\times \int_{-\infty}^{a} \int_{-\infty}^{b} dy\, dx\, e^{-\frac{s^2}{2}\left(\frac{q-1}{6-4q}\right)\frac{x^2 + y^2 - 2\rho xy}{1 - \rho^2}}$$

$$= \frac{2^{2 - \frac{1}{q-1}}(q-1)}{\Gamma\left(\frac{1}{q-1} - 1\right)(6 - 4q)} \int_0^{\infty} ds\, s^{\frac{2}{q-1} - 1} e^{-\frac{s^2}{2}} \Phi\left(s\sqrt{\frac{q-1}{6-4q}}[a, b], \rho\right). \tag{37}$$

$\Phi(\mathbf{x}, \rho)$ corresponds to the bivariate gaussian cumulative distribution function of $\mathbf{x}$ with correlation coefficient ρ. Then, by taking the derivative with respect to ρ we obtain

$$\frac{d\Phi_q([a, b], \rho)}{d\rho} = \frac{2^{2-\frac{1}{q-1}}(q-1)}{\Gamma\left(\frac{1}{q-1}-1\right)(6-4q)} \int_0^\infty ds\, s^{\frac{1}{q-1}} e^{-\frac{s^2}{2}} \frac{e^{-\frac{s^2}{2}\left(\frac{q-1}{6-4q}\right)\frac{a^2+b^2-2\rho ab}{1-\rho^2}}}{2\pi\sqrt{1-\rho^2}}. \tag{38}$$

With the change of variables $r = s\sqrt{\left[1+\left(\frac{q-1}{6-4q}\right)\frac{a^2+b^2-2\rho ab}{1-\rho^2}\right]}$,

$$\frac{d\Phi_q([a, b], \rho)}{d\rho} = \frac{2^{2-\frac{1}{q-1}}(q-1)}{\Gamma\left(\frac{1}{q-1}-1\right)(6-4q)} \frac{\left[1+\left(\frac{q-1}{6-4q}\right)\frac{a^2+b^2-2\rho ab}{1-\rho^2}\right]^{\frac{1}{1-q}}}{2\pi\sqrt{1-\rho^2}} \int_0^\infty dr\, r^{\frac{2}{q-1}-1} e^{-\frac{r^2}{2}}$$

$$= \frac{\Gamma\left(\frac{1}{q-1}\right)(q-1)}{\pi\Gamma\left(\frac{1}{q-1}-1\right)(6-4q)\sqrt{1-\rho^2}} \left[1+\left(\frac{q-1}{6-4q}\right)\frac{a^2+b^2-2\rho ab}{1-\rho^2}\right]^{\frac{1}{1-q}}, \tag{39}$$

since $\int_0^\infty dr\, r^{\frac{2}{q-1}-1} e^{-\frac{r^2}{2}} = 2^{\frac{1}{q-1}-1}\Gamma\left(\frac{1}{q-1}\right)$. This formula can be integrated with respect to ρ to produce

$$\Phi_q([a, b], \rho) = \Phi_q([a, b], s) + \frac{\Gamma\left(\frac{1}{q-1}\right)(q-1)}{\pi\Gamma\left(\frac{1}{q-1}\right)(6-4q)} \int_s^\rho dr \frac{\left[1+\left(\frac{q-1}{6-4q}\right)\frac{a^2+b^2-2rab}{1-r^2}\right]^{\frac{1}{1-q}}}{\sqrt{1-r^2}}, \tag{40}$$

where $s = sign(\rho)$ and so,

$$\Phi_q([a, b], s) = \begin{cases} \Phi_q(\min\{a, b\}) & \text{if } s = 1 \\ \max\{0, \Phi_q(\min\{a, b\}) - \Phi_q(\min\{a, b\})\} & \text{if } s = -1 \end{cases} \tag{41}$$

with $\Phi_q(\cdot)$ defined as the univariate q-Gaussian distribution. To implement Eq. (40) numerically in an algorithm, the change of variables $r = \sin(\theta)$ is desired. Thus, the bivariate q-Gaussian cumulative distribution function is calculated as

$$\Phi_q([a, b], \rho) = \Phi_q([a, b], s) + \frac{\Gamma\left(\frac{1}{q-1}\right)(q-1)}{\pi\Gamma\left(\frac{1}{q-1}\right)(6-4q)} \int_{s\,\pi/2}^{\sin^{-1}\rho} dr \left[1+\left(\frac{q-1}{6-4q}\right)\frac{a^2+b^2-2ab\sin\theta}{\cos^2\theta}\right]^{\frac{1}{1-q}}. \tag{42}$$

Note that the cumulative distribution function quantifies the amount of beyond-pairwise correlations in inputs through the parameter q.

3.3. Homogeneous population model

We provide in the following the derivation of the asymptotic spike count distribution for a large homogeneous (for analytical tractability) population of neurons for the q-DG model in terms of the Student's t-distribution, as alternative of the one previously presented in [34]. This provides a much easier implementation of the algorithms when quantifying HOCs.

We consider a pool of N neurons where each unit has a membrane potential u_i subject to a joint q-Normal distribution. Given the N-dimensional q-Gaussian random variable U, where $A = \mathbb{I}(1-\alpha) + \alpha\mathbb{1}_N\mathbb{1}_N^T$, then

$$u_i = h + \sqrt{1-\alpha}\,v_i + \sqrt{\alpha}\varepsilon. \tag{43}$$

The variables v_i are independent random variables subject to N_q and ε corresponds to Gaussian noise. Then, the u_i are independent for a given ε. Within the q-DG model, the binary patterns X are generated by $X = 1$ (see Section 2.1). The input statistics are chosen such that the outputs X have mean μ and covariance Σ. Since we here focus on homogeneous populations then $\mu_i = \mu$ and $\Sigma_{ij} = \sigma$, and we take the pairwise correlation coefficient as

$$\rho = \frac{\sigma}{\mu(1-\mu)}. \tag{44}$$

By symmetry, all activity patterns with the same number of spikes are equally likely. Such a homogeneous population model is fully characterized by the number of neurons that spike synchronously, i.e. the population spike count distribution k. In this sense, in a population of N neurons, we can define $r = k/N$ as the proportion of neurons spiking at any time. Let us

now consider the probability of exactly $k = N \cdot r$ neurons firing within a given time window ΔT across a population of N neurons as

$$P\{k\} = Prob\{x_1 = x_2 = \cdots = x_k = 1, x_{k+1} = \cdots = x_N = 0\}, \tag{45}$$

where the spiking of neuron x_i is constrained by their membrane potential as previously explained. Note that in the Eq. (45) we have presented the compact form of the probability, however all possible combinations are taken into account to compute the probability as in [33,34]. The probability of having a certain pattern with k spikes is given by the expectation value taken with respect to the random variable ε, and $P\{\cdot|\varepsilon\}$ is the conditional probability for ε. This allows us to calculate the probability of having $r = k \cdot N$ neurons firing, separating the contribution of neurons that are spiking $P\{u > 0|\varepsilon\}^k$ from those that are silent $P\{u \le 0|\varepsilon\}^{N-k}$, as

$$P\{k\} = E_\varepsilon \left[P\{u_1; \cdots ; u_k > 0, u_{k+1}; \cdots ; u_N \le 0|\varepsilon\} \right]$$
$$\equiv E_\varepsilon \left[\binom{N}{k} (P(u > 0|\varepsilon))^k (P(u \le 0|\varepsilon))^{N-k} \right]. \tag{46}$$

Then, for each neuron,

$$P(u_i > 0) = P\left(v_i < \frac{\sqrt{\alpha}\varepsilon + h}{\sqrt{1-\alpha}} \right)$$
$$= \Phi_q \left(\frac{\sqrt{\alpha}\varepsilon + h}{\sqrt{1-\alpha}} \right)$$
$$\equiv \frac{1}{\zeta} T_\nu \left(\frac{\sqrt{\alpha}\varepsilon + h}{\zeta\sqrt{1-\alpha}} \right) = F_q(\varepsilon). \tag{47}$$

Here, T_ν is the cumulative distribution function of a Student's t-distribution with ν degrees of freedom, and ζ is the scale parameter. The asymptotic rate distribution for the homogeneous population, using the saddle–node approximation [33,34, 38,40,61], can be then derived. Under this approximation, the joint distribution of firing can be estimated as

$$f_q(r) \cong N P\{r = k/N\}$$
$$= N E_\varepsilon \left[\binom{N}{k} (F_q(\varepsilon))^k (1 - F_q(\varepsilon))^{N-k} \right]$$
$$= \sqrt{\frac{1}{2\pi r(1-r) \mid z_q''(\varepsilon_0) \mid}} \exp\left(N z_q(\varepsilon_0) - \frac{\varepsilon_0}{2} \right), \tag{48}$$

where $z_q(\varepsilon) = r \log\left(\frac{F_q(\varepsilon)}{r} \right) + (1-r) \log\left(\frac{1-F_q(\varepsilon)}{1-r} \right)$, and ε_0 the value of ε that maximizes z_q. That is, $\varepsilon_0 = F_q^{-1}(r)$, which implies that $r = F_q(\varepsilon_0)$, where r can take values between [0, 1] and ε_0 is defined for all real numbers. The first derivative of $z_q(\varepsilon)$ with respect to ε yields

$$z_q'(\varepsilon) = \frac{r}{F_q(\varepsilon)} - \frac{1-r}{1-F_q(\varepsilon)}, \tag{49}$$

and hence the second derivative is

$$z_q''(\varepsilon) = -\left[\frac{r}{(F_q(\epsilon))^2} + \frac{1-r}{(1-F_q(\epsilon))^2} \right] (F_q'(\varepsilon))^2 + \left(\frac{r}{F_q(\varepsilon)} - \frac{1-r}{1-F_q(\varepsilon)} \right) F_q''(\varepsilon). \tag{50}$$

As we are working within the saddle-point approximation, the parameter ε_0 maximizes $z_q(\varepsilon)$, and thus [33],

$$\varepsilon = \varepsilon_0 \Longrightarrow z_q'(\varepsilon_0) = 0 \wedge F_q(\varepsilon_0) = r. \tag{51}$$

Then,

$$z_q''(\varepsilon_0) = -\left(\frac{1}{r} - \frac{1}{1-r} \right) (F_q'(\varepsilon_0))^2. \tag{52}$$

Considering Eq. (47), we can easily calculate previous quantities in terms of the cumulative distribution function of a Student's t-distribution since its derivative is well known. Therefore, the probability density for the homogeneous q-DG distribution is

$$f_\nu(r) = \zeta \sqrt{\frac{1-\alpha}{\alpha}} \left[1 + \frac{(T_\nu^{-1}(\zeta r))^2}{\nu} \right]^{-\frac{\nu+1}{2}} \exp\left[-\frac{\zeta}{2} \frac{1-\alpha}{\alpha} \left(T_\nu^{-1}(\zeta r) - \frac{h}{\zeta\sqrt{1-\alpha}} \right)^2 \right], \tag{53}$$
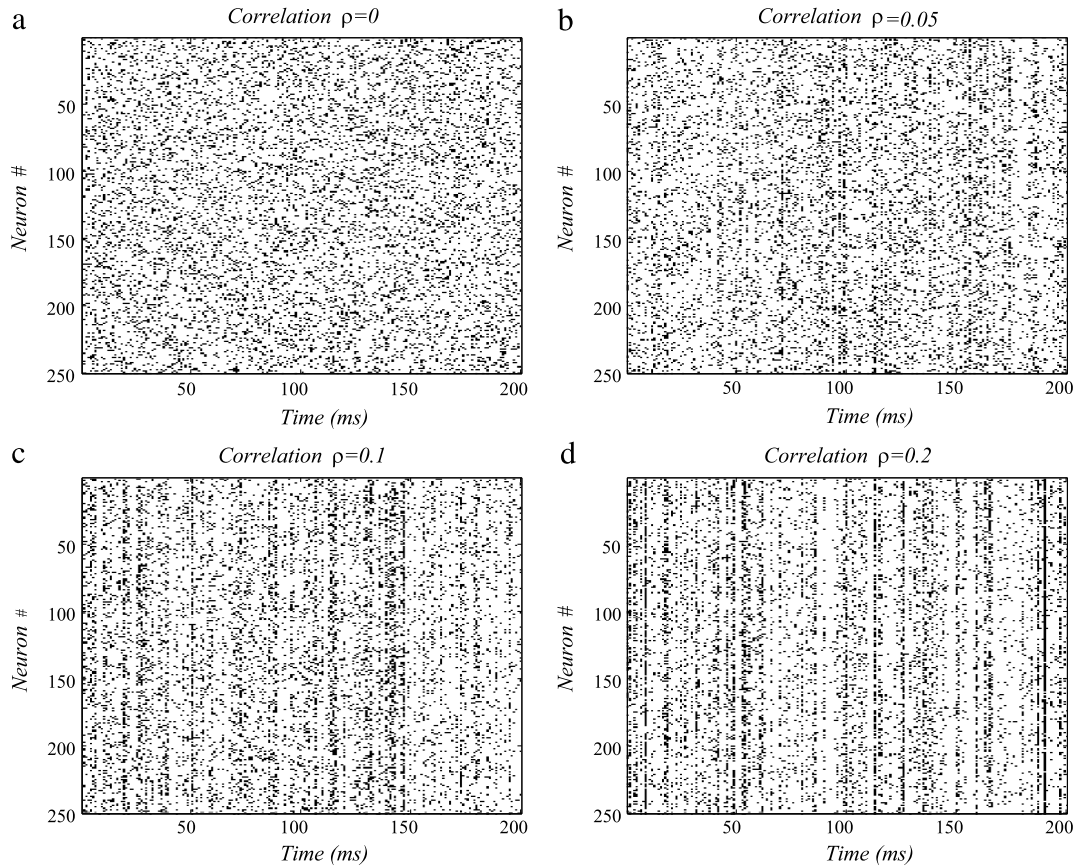
Fig. 1. Raster plots of synthetically generated multi-neuron firing patterns within the DG formalism ($q = 1$) and considering different output correlations ρ. (a) $\rho = 0$, (b) $\rho = 0.05$, (c) $\rho = 0.1$ and (d) $\rho = 0.2$.

or, in terms of the deformation parameter q, it can be written as

$$
f_q(r) = \sqrt{\frac{(1-\alpha)(5-3q)}{\alpha(3-q)}}\, \exp_q^{-1}\left[-\left(\frac{q-1}{3-q}\right)\left(\mathrm{T}^{-1}_{\left(\frac{3-q}{q-1}\right)}\left(\sqrt{\frac{5-3q}{3-q}}r\right)\right)^2\right]
$$

$$
\times \exp\left[-\frac{\sqrt{5-3q}(1-\alpha)}{2\alpha(3-q)}\left(\mathrm{T}^{-1}_{\left(\frac{3-q}{q-1}\right)}\left(\sqrt{\frac{5-3q}{3-q}}r\right) - \frac{h\sqrt{3-q}}{\sqrt{(5-3q)(1-\alpha)}}\right)^2\right].
\tag{54}
$$

3.4. Numerical results

In the following, we present a careful comparison of our current method with the DG model [24,38–40]. This is important to investigate how the quantification of HOCs in the neuronal inputs of the q-DG model differs from the DG model [24,39,40]. Quantifying the degree of HOCs through the parameter q can help to further understand how information is processed in cortical networks, and to obtain a deeper knowledge of the nonlinear spiking dynamics within a neuronal ensemble. In the DG model, correlations between neurons arise from Gaussian inputs into threshold neurons, and therefore have no interactions beyond second order [24,39,40]. In contrast, in our model correlations between neurons may also arise from non-Gaussian inputs into threshold neurons (Gaussian inputs are recovered by taking the limit of the deformation parameter $q \to 1$).

First, we consider the asymptotic regime to investigate the effect of HOCs in the common inputs and the spiking outputs of a population of neurons. Figs. 1 and 2(a)–(d) show the raster plots when considering $q = 1$ (DG model) and in the case with a non-Gaussian distributions with a deformation parameter of $q = 1.35$, respectively, for $\rho = 0$, $\rho = 0.05$, $\rho = 0.1$ and $\rho = 0.2$. We are considering beyond-pairwise correlations when the values of the deformation parameter q are larger than one. This leads to a higher amount of synchronization among spikes and the effect becomes more pronounced as ρ increases. That is, having $q = 1.35$ induces a more considerable amount of synchronized spikes, as it can be observed by
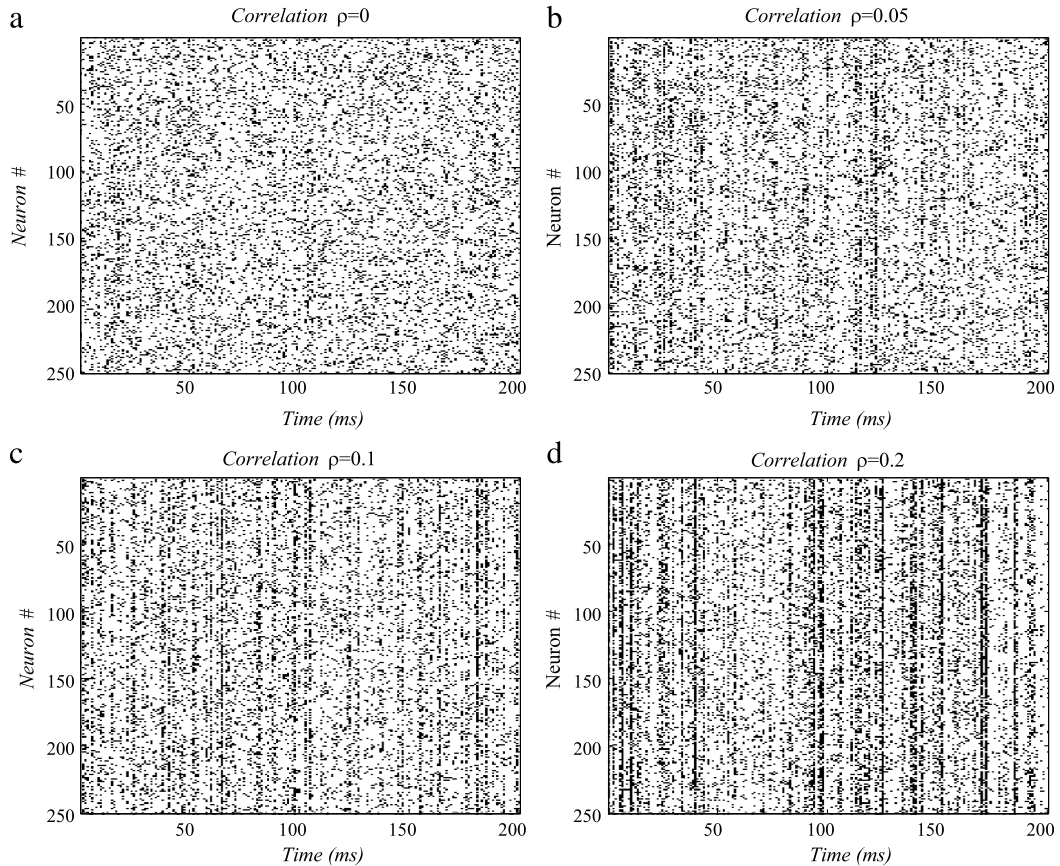
Fig. 2. Same as in Fig. 1 but considering *q*-DG formalism with $q = 1.35$. (a) $\rho = 0$, (b) $\rho = 0.05$, (c) $\rho = 0.1$ and (d) $\rho = 0.2$. Note that in comparison with Fig. 1, higher values of spike synchronization are visible at simple sight due to the larger amount of input correlation.

a simple comparison between the raster plots of Figs. 1 and 2. The latter provide a first insight of how input correlations may affect the spiking output activity. However, a more accurate analysis is needed to precisely quantify the effect of HOCs. Fig. 3(a) and (b) show the joint distributions of firing Eq. (54) when considering a Gaussian distribution ($q = 1$) and a non-Gaussian distribution with $q = 1.35$, respectively, taking $\rho = 0$, $\rho = 0.05$, $\rho = 0.1$ and $\rho = 0.2$. As expected, a larger deformation parameter prompts larger excursions of synchronized spikes [33,34]. Second, we precisely analyze how HOCs affect the output correlations. More specifically, we quantify how correlations between the discharges of neurons are affected by HOCs between their inputs, and whether they could produce deviations from pure pairwise modeling (i.e. the DG model). In order to do so, we provide a detailed analysis of how the output spike correlations deviate from the DG model as the deformation parameter *q* increases when considering different values of mean and variance in the neuronal inputs. Eq. (24) elicits a characteristic relationship between the pairwise correlations of the inputs and the pairwise correlations of the binary threshold neurons. That is, given the same amount of input correlation α, the correlation of the output ρ will grow monotonically with their firing rates. Fig. 4(a), (b) and (c) show the dependence of the output correlation ρ versus input correlation α, when considering different values of the firing rate $\mu = 0.1$, $\mu = 0.25$ and $\mu = 0.5$, respectively, for $q = 1$, $q = 1.1$, $q = 1.2$ and $q = 1.3$. Note that increasing the amount of HOCs in the input connectivity leads to slightly lower values of synchronized spikes in the neural population output. However, larger values of the firing rate increase the amount synchrony.

In order to gain a better understanding of how cortical cells may transform correlation between their synaptic inputs currents into correlation between their output spike trains, we also need to investigate the output correlation ρ as function of the firing rate μ. Eq. (24) also provides a means for obtaining this relationship when input correlations are kept constant. Fig. 5(a), (b), and (c) show how the output correlation ρ varies as function of the firing rate μ, for different values of the input correlation $\alpha = 0.1$, $\alpha = 0.2$ and $\alpha = 0.3$, respectively, for $q = 1$, $q = 1.1$, $q = 1.2$ and $q = 1.3$. Importantly, as it can appreciated from the previous figure, HOCs in the inputs yield a lower amount synchrony when considering small firing rates ($\mu < 0.15$). When $\mu = 0.15$ the output correlation is equal to the one that is obtained in the DG model. In contrast, when the firing rate value becomes higher ($\mu > 0.15$), that leads to larger excursions of synchronized spikes as the deformation parameter *q* grows. The relationship between output correlation and firing rate is sensitive to input heterogeneities and, even when the higher-order input correlation remains fixed, the output correlation increases with the firing rate. However,
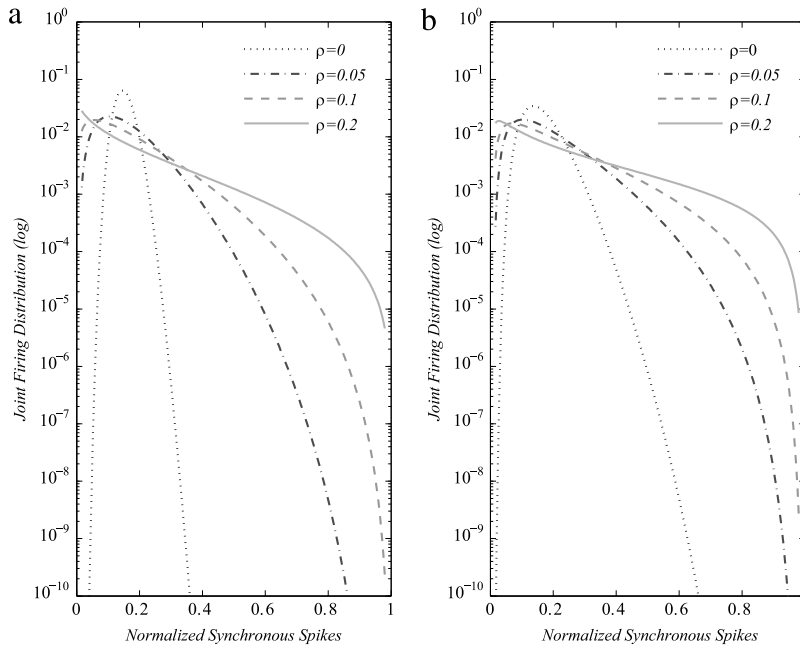
Fig. 3. Joint distribution of firing as function of the normalized number of spikes (semi-log in the Y axis). (a) Considering $q = 1$ and $\mu = 15$ for $\rho = 0$, $\rho = 0.05$, $\rho = 0.1$ and $\rho = 0.2$. (b) Considering $q = 1.35$ and $\mu = 0.15$ for $\rho = 0$, $\rho = 0.05$, $\rho = 0.1$ and $\rho = 0.2$.

it displays a complex behavior as the order of correlation in their inputs becomes higher than two. That is, HOCs induce more synchronous silent firing when $\mu < 0.15$, while for higher values of μ the degree of synchrony grows nonlinearly.

Finally, we want to determine the effect of HOCs in common input on the entropy of the population. Fig. 6 shows that the degree of chaos results strongly dependent on HOCs. That is, Fig. 6 illustrates that the porcentual increment of entropy grows with the parameter q, using Eq. (27). Furthermore, in Fig. 7 we show how statically different, using Eq. (29), the joint distribution of firing is when considering HOCs. Note that the relative KL-divergence grows as the degree of correlation becomes higher, quantifying the inefficiency of incorrectly assuming that the distribution is the one that corresponds to $f_1(\rho)$ when the true distribution is $f_q(\rho)$. Thus, the amount of information in the spiking output of a neural population cannot be computed without knowing the correlational structure in their inputs, which is quantified by the deformation parameter q.

4. Discussion and conclusions

As we have discussed in the previous sections, within the DG model common inputs have no interactions beyond second order [38–40]. More specifically, in the DG model HOCs in the spiking output are evaluated by investigating how the probability distribution generated by this model differs from the distribution one would obtain with a pure pairwise model (i.e. Ising). In order to do so, different means and variances in the common inputs are considered [38–40]. Thus, despite that the DG model captures several features of beyond-pairwise correlations in the spiking output, it does not explicitly quantify them. It just captures HOCs in the population firing distribution that cannot be produced by pure pairwise models [38–40,24]. Inputs are modeled by correlated Gaussians, with mean h and covariance A chosen such that the output X has mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

In this paper, we present a natural extension of the DG model by including HOCs within the neuronal input. Importantly, we also quantify the amount of synchronization in the spiking outputs thought ρ and the deformation parameter q. That is, our model naturally quantifies the degree of HOCs in the common neuronal input and in the spiking output. In order to efficiently account for the complex structure of HOCs in a neural population, we build a model that considers all nonlinearities present through the deformation parameter q. We show that, in comparison with the DG model, in the current approach the input is modeled by a correlated q-Gaussian with mean h and variance A, which are chosen such that the output X has mean $\boldsymbol{\mu}$ and pairwise variance $\boldsymbol{\Sigma}$. This allows us to precisely quantify the amount of correlations that are generated from higher-order interconnectivity of the common overlapping neuronal inputs throughout the deformation parameter q. Moreover, output pairwise and higher-order spike synchronization in the neuronal ensemble is quantified by ρ and q, respectively. It is important to remark on the three major differences between the approaches we developed in [33,34] and our current theoretical model: first in this paper we explicitly provide an exhaustive analysis of how input statistics are transformed in this thresholding process into output statistics; second we supply an efficient formalism and numerical implementation for the model accounting for HOCs in the neural population inputs and outputs, and third, we show under
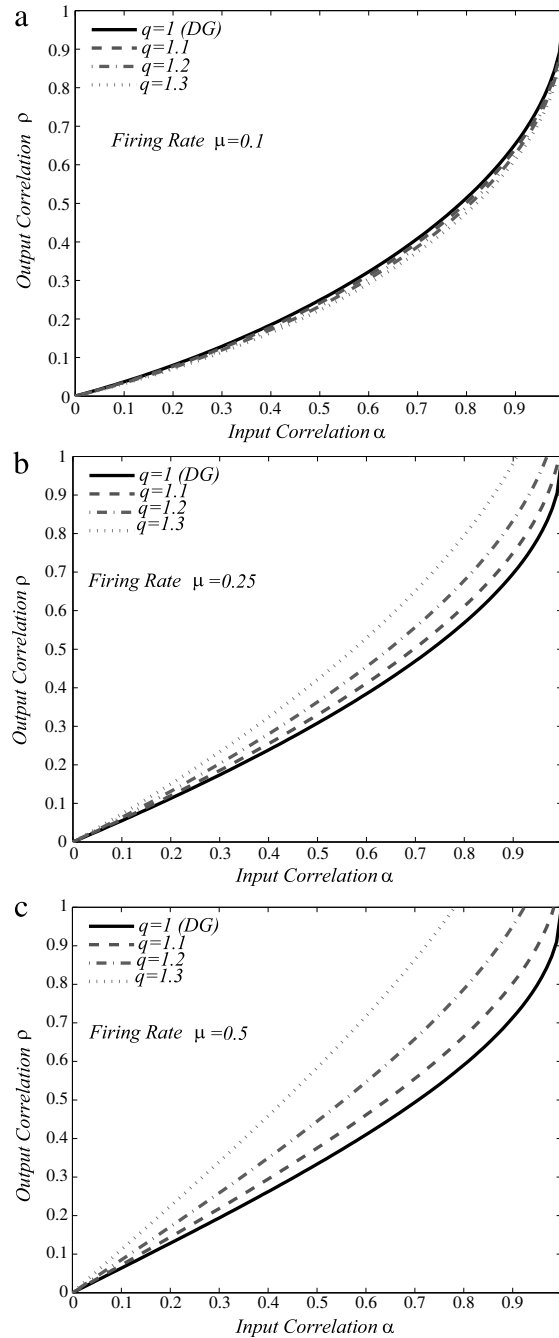
Fig. 4. Output correlation ρ versus the input correlation α, taking $q = 1$, $q = 1.1$, $q = 1.2$ and $q = 1.3$, when μ is fixed. (a) Considering $\mu = 0.1$. (b) $\mu = 0.25$. (c) $\mu = 0.5$. Note that higher values of q increases the degree of the output synchronization ρ as the firing rate μ becomes larger.

which conditions HOCs in the common inputs can lead to either bigger or smaller number of synchronized spikes in the neural population output.

Summarizing we develop a model, in which correlations across neurons arise from q-Gaussian inputs into threshold neurons generating HOC spiking outputs. It is therefore an extension of the DG model proposed by Amari [38–40], where the inputs to the model are Gaussian distributed and thus have no interactions beyond second order in their input. Our current theoretical formalism relies on recent progress made on the ECLT, and using mathematical tools of non-extensive statistical mechanics [46,49–53,56,62,63], we provide a flexible framework for efficient simulation and modeling of correlated spike trains with specified HOCs, even if their origin and source are unknown. Furthermore, it can be implemented numerically
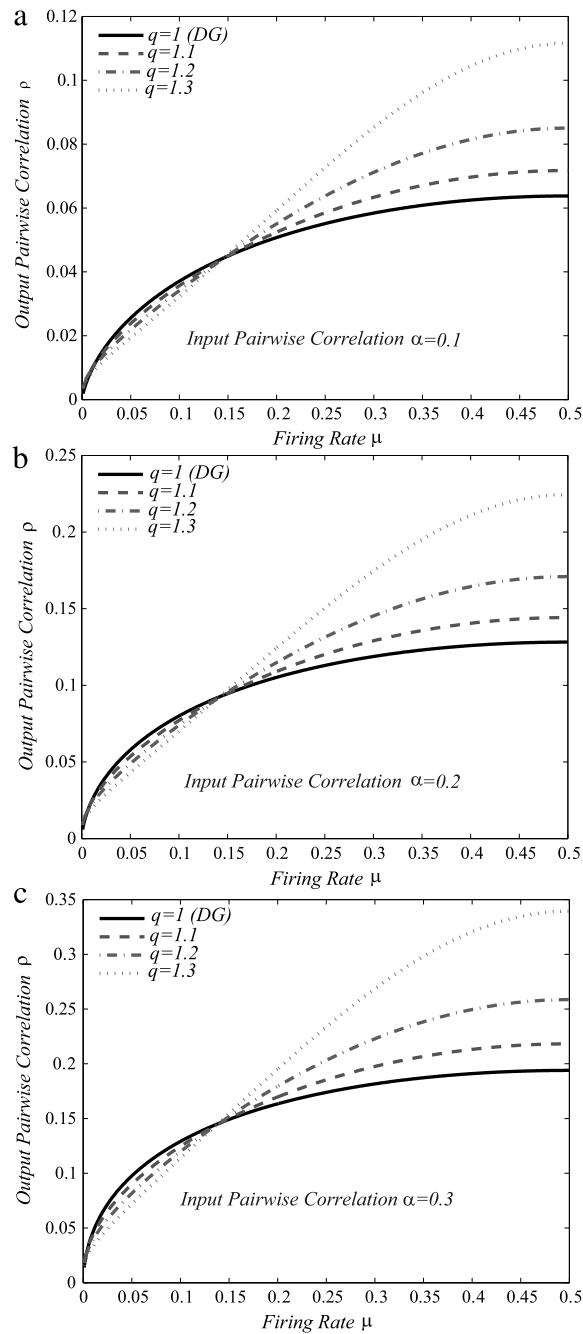
Fig. 5. Output correlation ρ as function of the firing rate μ, taking $q = 1$, $q = 1.1$, $q = 1.2$ and $q = 1.3$. (a) Considering a fixed input correlation $\alpha = 0.1$. (b) Input correlation $\alpha = 0.2$. (c) Input correlation $\alpha = 0.3$.

with ease as we expressed the joint probability distribution of firing in terms of the widely used Student's t-distribution. Moreover, in our model, not only a change in the mean input modifies interactions of all orders but also the amount of HOCs q does, and therefore it can help us gain further understanding of the role of higher-order interactions in stimulus coding. The current approach captures higher-order redundancies which increase as the q parameter becomes larger, leading to a higher amount of randomness as the entropy grows and thus an increased amount of chaos. Importantly, through the divergent KL-divergence we quantify the inefficiency of incorrectly assuming that the distribution is made by the DG model showing that HOCs have an important effect at population level. Given the ubiquity of common input in sensory systems, the model thus gives quantitative predictions for the conditions under which HOCs will be important, quantifying the inefficiency of incorrectly assuming that neurons have just pairwise correlations in their common inputs.
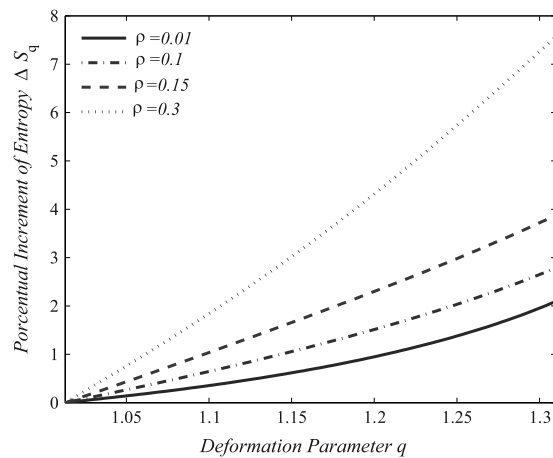
Fig. 6. Porcentual increment of entropy as function of the deformation parameter q as in Eq. (27) considering different output correlation values $\rho = 0.01$, $\rho = 0.1$, $\rho = 0.15$ and $\rho = 0.3$. Note that the amount chaos increases as higher-order input correlations become more relevant.
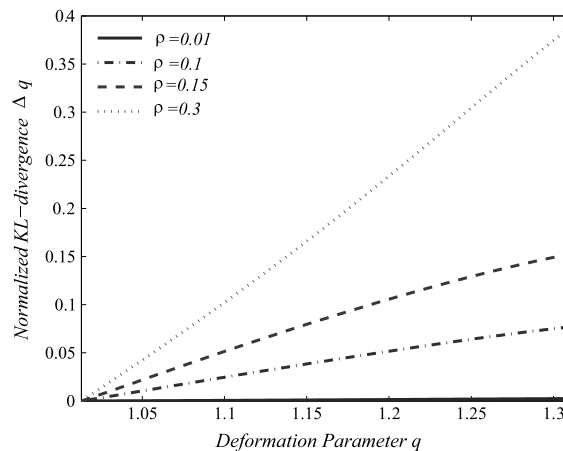

Fig. 7. The normalized Kullback–Leibler divergence versus the deformation parameter q, given different output correlation values $\rho = 0.01$, $\rho = 0.1$, $\rho = 0.15$ and $\rho = 0.3$ (the reference distribution is taken as $\rho = 0$ for each value of ρ, and it is normalized to the distribution with $q = 1$ as in Eq. (29)).

The current formalism also provides a robust and conceptually adequate framework for generating synthetic spike trains with a wide range of different firing rates and HOCs' structures. Generation of spike trains by binarizing a multivariate q-Gaussian random variable is possible even for large populations of neurons and for spike count statistics with arbitrary marginal distributions. This provides an efficient methodology to generate spike trains with an specified higher-order correlation structure. Our technique can be applied to experiments in neurophysiology including microstimulation of populations of neurons, opening up the possibility of investigating natural stimulations that might generate with different correlation structures and to study their effect in driving neural responses.

Acknowledgment

References

[1] G. Werner, V.B. Mountcastle, Neural activity in mechanoreceptive cutaneous afferents: stimulus–response relations, weber functions and information transmission, J. Neurophysiol. 28 (1905) 359–397.
[2] D. Tolhurst, The amount of information transmitted about contrast by neurones in the cat's visual cortex, Vis. Neurosci. 2 (4) (1989) 409–413.
[3] S. Panzeri, S.R. Schultz, A unified approach to the study of temporal, correlational, and rate coding, Neural Comput. 13 (6) (2001) 1311–1349.
[4] S.R. Schultz, S. Panzeri, Temporal correlations and neural spike train entropy, Phys. Rev. Lett. 86 (25) (2001) 5823(4).
[5] A.K. Kreiter, W. Singer, Oscillatory neuronal responses in the visual cortex of the awake macaque monkey, Eur. J. Neurosci. 4 (4) (1992) 369–375.
[6] A.K. Kreiter, W. Singer, Stimulus-dependent synchronization of neuronal responses in the visual cortex of the awake macaque monkey, J. Neurosci. 16 (7) (1996) 2381–2396.
[7] R.C.D. Charms, M.M. Merzenich, Primary cortical representation of sounds by the coordination of action-potential timing, Nature 381 (1996) 610–613.

[8] T.J. Gawne, T.W. Kjaer, J.A. Herz, B.J. Richmond, Adjacent visual cortical complex cells share about 20% of their stimulus-related information, Cereb. Cortex 6 (3) (1996) 482–489.
[9] P.R. Roelfsema, A.K. Engel, P. Koenig, W. Singer, Visuomotor integration is associated with zero time-lag synchronization among cortical areas, Nature 385 (6612) (1997) 157–161.
[10] A. Kohn, M.A. Smith, Stimulus dependence of neuronal correlation in primary visual cortex of the macaque, J. Neurosci. 25 (14) (2005) 3661–3673.
[11] F. Montani, A. Kohn, M. Smith, S. Schultz, The role of correlations in direction and contrast coding in the primary visual cortex, J. Neurosci. 27 (9) (2007) 2338–2348.
[12] F. Montani, E.B. Deleglise, O.A. Rosso, Efficiency characterization of a large neuronal network: a causal information approach, Physica A 401 (2014) 58–70.
[13] F. Montani, O.A. Rosso, F. Matias, S.L. Bressler, C.R. Mirasso, A symbolic information approach to determine anticipated and delayed synchronization in neuronal circuit models, Philos. Trans. R. Soc. Lond. Ser. A 373 (2015) 20150110.
[14] F. Montani, R. Baravalle, L. Montangie, O. Rosso, Causal information quantification of prominent dynamical features of biological neurons., Philos. Trans. R. Soc. Lond. Ser. A 373 (2015) 20150109.
[15] F. Montani, A. Oliynyk, L. Fadiga, Superlinear summation of information in premotor neuron pairs, Int. J. Neur. Syst. 27 (2) (2017) 1650009.
[16] R. Moreno-Bote, J. Beck, I. Kanitscheider, X. Pitkow, P. Latham, A. Pouget, Nature Neurosci. 17 (2014) 1410–1417.
[17] A. Kohn, R. Coen-Cagli, I. Kanitscheider, A. Pouget, Correlations and neuronal population information, Annu. Rev. Neurosci. 39 (2016) 237–256.
[18] E. Schneidman, M.J. Berry, R. Segev, W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population, Nature 440 (2006) 1007–1012.
[19] J. Shlens, G.D. Field, J.L. Gauthier, M.I. Grivich, D. Petrusca, A. Sher, A.M. Litke, E.J. Chichilnisky, The structure of multi-neuron firing patterns in the primate retina, J. Neurosci. 26 (2006) 8254–8266.
[20] J.M. Alonso, C.-I. Yeh, C.R. Stoelzel, Visual stimuli modulate precise synchronous firing within the thalamus, Thalamus Relat. Sys. 4 (2008) 21–34.
[21] J. Samonds, A. Bonds, Gamma oscillation maintains stimulus structure dependent synchronization in cat visual cortex, J. Neurophysiol. 93 (2005) 223–236.
[22] S.R. Schultz, K. Kitamura, J. Krupic, A. Post-Uiterweer, M. Hausser, Spatial pattern coding of sensory information by climbing-fiber evoked calcium signals in networks of neighboring cerebellar purkinje cells, J. Neurosci. 29 (2009) 8005–8015.
[23] A.K. Wise, N.L. Cerminara, D.E. Marple-Horvat, R. Apps, Mechanisms of synchronous activity in cerebellar purkinje cells, J. Physiol. 588 (2010) 2373–2390.
[24] S. Yu, H. Yang, H. Nakahara, G. Santos, D. Nikolic, D. Plenz, Higher-order interactions characterized in cortical activity, J. Neurosci. 31 (48) (2011) 17514–17526.
[25] G. Tkacik, E. Schneidman, W. Bialek II, W. Bialek, Ising models for networks of real neurons., qbio.NC/0611072.
[26] J. Shlens, G.D. Field, J.L. Gauthier, M. Greschner, A. Sher, A. Litke, E. Chichilnisky, The structure of large-scale synchronized firing in primate retina, J. Neurosci. 29 (2009) 5022–5031.
[27] P. Berens, M. Bethge, Near-maximum entropy models for binary neural representations of natural images, in: Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems, 20, MIT Press, Cambridge, MA, 2008, pp. 97–104.
[28] Y. Roudi, S. Nirenberg, P. Latham, Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't, PLoS Comput. Biol. 5 (2009) e1000380.
[29] I. Ohiorhenuan, J. Victor, Information-geometric measure of 3-neuron firing patterns characterizes scale-dependence in cortical networks, J. Comput. Neurosci. 30 (2010) 125–141.
[30] I. Ohiorhenuan, K. Mechler, F. Purpura, A. Schmid, Q. Hu, J. Victor, Sparse coding and high-order correlations in fine-scale cortical networks, Nature 466 (2010) 617–622.
[31] E. Ganmor, R. Segev, E. Schneidman, Sparse low-order interaction network underlies a highly correlated and learnable neural population code, Proc. Natl. Acad. Sci. 108 (2011) 9679–9684.
[32] F. Montani, R. Ince, R. Senatore, E. Arabzadeh, M. Diamond, S. Panzeri, The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex, Phil. Trans. R. Soc. A 367 (2009) 3297–3310.
[33] F. Montani, E. Phoka, M. Portesi, S.R. Schultz, Statistical modelling of higher-order correlations in pools of neural activity, Physica A 392 (2013) 3066–3086.
[34] L. Montangie, F. Montani, Quantifying higher-order correlations in a neuronal pool, Physica A 421 (2015) 388–400.
[35] L. Montangie, F. Montani, Effect of interacting second- and third-order stimulus-dependent correlations on population-coding asymmetries, Phys. Rev. E 94 (2016) 042303.
[36] N.A. Cayco-Gajic, J. Zylberberg, E. Shea-Brown, Triplet correlations among similarly tuned cells impact population coding, Front. Comput. Neurosci. 9 (2015) 57.
[37] H. Shimazaki, S.I. Amari, E.N. Brown, S. Grun, State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data, PLoS Comput. Biol. 8 (2012) e1002385.
[38] S. Amari, H. Nakahara, S. Wu, Y. Sakai, Synchronous firing and higher-order interactions in neuron pool, Neural Comput. 15 (2003) 127–142.
[39] J.H. Macke, P. Berens, A.S. Ecker, A.S. Tolias, M. Bethge, Generating spike trains with specified correlation coefficients, Neural Comput. 21 (2009) 397–423.
[40] J.H. Macke, M. Opper, M. Bethge, Common input explains higher order correlations and entropy in a simple model of neural population activity, Phys. Rev. Lett. 106 (2011) 208102.
[41] D.R. Cox, N. Wermuth, On some models for multivariate binary variables parallel in complexity with the multivariate gaussian distribution, Biometrika 89 (2) (2002) 462–469.
[42] J. de la Rocha, B. Doiron, E. Shea-Brown, K. Josic, A. Reyes, Correlation between neural spike trains increases with firing rate, Nature 448 (7155) (2007) 802–806.
[43] D.R. Lyamzin, J.H. Macke, N.A. Lesica, Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations, Front. Comp. Neurosci. 4 (2010) 144.
[44] M.R. DeWeese, A.M. Zador, Non-gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex, J. Neurosci. 26 (47) (2006) 12206–12218.
[45] A. Roxin, N. Brunel, D. Hansel, G. Mongillo, C. van Vreeswijk, On the distribution of firing rates in networks of cortical neurons, J. Neurosci. 31 (45) (2011) 16217–16226.
[46] S. Umarov, C. Tsallis, S. Steinberg, On a q-central limit theorem consistent with nonextensive statistical mechanics, Milan J. Math. 76 (2008) 307–328.
[47] C. Tsallis, Possible generalization of boltzmann-gibbs statistics, J. Stat. Phys. 52 (1988) 479–487.
[48] S.M. Ross, A First Course in Probability, Vol, second ed., Macmillan Publishing Company, 1976.
[49] S. Umarov, C. Tsallis, M. Gell-Mann, S. Steinberg, Generalization of symmetric α-stable Lévy distributions for $q > 1$, J. Math. Phys. 51 (2010) 033502.
[50] C. Vignat, A. Plastino, Central limit theorem and deformed exponentials, J. Phys. A: Math. Theor. 40 (2007) 969–978.
[51] C. Vignat, A. Plastino, Estimation in a fluctuating medium and power-law distributions, Phys. Lett. A 360 (2007) 415–418.
[52] C. Vignat, A. Plastino, Scale invariance and related properties of q-gaussian systems, Phys. Lett. A 365 (2007) 370–375.
[53] C. Tsallis, in: J.L. Morán-Lopez, J.M. Sánchez (Eds.), New Trends in Magnetism, Magnetic Materials and Their Applications, Plenum, New York, 1994, p. 451.
[54] S. Amari, A. Ohara, Geometry of q-exponential family of probability distributions, Entropy 13 (2011) 1170–1185.
[55] S. Nakahara, S. Amari, Information geometric measure for neural spikes, Neural Comput. 14 (2002) 2269–2316.
[56] S. Amari, H. Nagaoka, Methods of Information Geometry, in: Translations of Mathematical Monographs, vol. 191, Oxford University Press, 2000.
[57] Z.F.H. Block, A multivariate extension of hoeffding lemma, Ann. Probab. 16 (4) (1988) 1803–1820.

[58] C. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Champaign, IL, USA, 1949.
[59] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Statist. 22 (1) (1951) 79–86.
[60] I.S. Gradsteyn, I. Ryzhik, Table of Integrals, Series and Products, seventh ed., Academic Press, 2007.
[61] R. Butler, Saddlepoint Approximations with Applications, in: Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2007.
[62] C. Tsallis, M. Gell-Mann, Y. Sato, Asymptotically scale-invariant occupancy of phase space makes the entropy s_q extensive, Proc. Natl. Acad. Sci. 102 (2000) 15377–15382.
[63] M.P.H. Stumpf, M.A. Porter, Critical truths about power laws, Science 335 (2012) 665–666.